


 Cite this: *RSC Adv.*, 2021, **11**, 17478

## Computational estimation of potential inhibitors from known drugs against the main protease of SARS-CoV-2†

 Nguyen Minh Tam, <sup>\*ab</sup> Minh Quan Pham, <sup>cd</sup> Nguyen Xuan Ha, <sup>e</sup>  
 Pham Cam Nam <sup>f</sup> and Huong Thi Thu Phung <sup>g\*</sup>

The coronavirus disease (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has rapidly spread worldwide recently, leading to global social and economic disruption. Although the emergently approved vaccine programs against SARS-CoV-2 have been rolled out globally, the number of COVID-19 daily cases and deaths has remained significantly high. Here, we attempt to computationally screen for possible medications for COVID-19 *via* rapidly estimating the highly potential inhibitors from an FDA-approved drug database against the main protease (Mpro) of SARS-CoV-2. The approach combined molecular docking and fast pulling of ligand (FPL) simulations that were demonstrated to be accurate and suitable for quick prediction of SARS-CoV-2 Mpro inhibitors. The results suggested that twenty-seven compounds were capable of strongly associating with SARS-CoV-2 Mpro. Among them, the seven top leads are daclatasvir, teniposide, etoposide, levoleucovorin, naldemedine, cabozantinib, and irinotecan. The potential application of these drugs in COVID-19 therapy has thus been discussed.

Received 31st March 2021

Accepted 3rd May 2021

DOI: 10.1039/d1ra02529e

[rsc.li/rsc-advances](http://rsc.li/rsc-advances)

### Introduction

The ongoing deadly pneumonia disease (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has uncontrollably spread globally since its first identification in December 2019. In March 2020, the WHO (World Health Organization) classified the COVID-19 outbreak as a “Global Pandemic”.<sup>1</sup> The transmission rate of the viral infection is extremely high<sup>2</sup> while the fatality rate ranges from 1% to 12%.<sup>3</sup> Recently, the COVID-19 disease has affected 221 countries and territories around the world with more than 153 million reported cases and 3.2 million deaths.<sup>4</sup>

SARS-CoV-2 has a single-stranded positive-sense RNA genome with a length of approximately 29.9 kb.<sup>5,6</sup> The viral genome is composed of 11 open reading frames coding for more than 20 different proteins. Among them, the SARS-CoV-2 main protease (Mpro), one of the most important proteins during the viral translation, is required to digest polyproteins at eleven or more conserved cleavage sites to produce various functional proteins.<sup>7</sup> The polypeptides generated are critical for the viral transcription and replication during its infection. The SARS-CoV-2 Mpro is well conserved in the Coronaviridae family<sup>8</sup> while its closely related homologs in humans have lacked. These characteristics cause Mpro one of the most interesting targets for the selection of antiviral drugs to restrain the SARS-CoV-2 growth and replication.<sup>9,10</sup> Consequently, multiple investigations have been done to define the promising inhibitors of this protease.<sup>7,11–18</sup>

Nowadays, the computer-aided drug design (CADD) approach has been broadly shown to remarkably save time and cost in the development of a new medication.<sup>19,20</sup> In CADD, the ligand-binding free energy  $\Delta G$  of an inhibitor with the targeted protein can be accurately predicted *via* molecular dynamics (MD) simulations.<sup>19,20</sup> Accordingly, rapidly and precisely estimating the ligand-binding free energy is extremely essential for identifying potential inhibitors.<sup>21,22</sup> Previously, the fast pulling of ligand (FPL) simulations was shown to effectively and accurately estimate the relative binding affinities of small molecules against HIV-1 (human immunodeficiency virus 1) protease or CHK1 (checkpoint kinase 1) with an affordable Central Processing Unit

<sup>a</sup>Computational Chemistry Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam. E-mail: [nguyenminhtam@tdtu.edu.vn](mailto:nguyenminhtam@tdtu.edu.vn)

<sup>b</sup>Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam

<sup>c</sup>Institute of Natural Products Chemistry, Vietnam Academy of Science and Technology, Hanoi, Vietnam

<sup>d</sup>Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

<sup>e</sup>Faculty of Chemistry and Environment, Thuyloi University, Ministry of Agriculture and Rural Development, Hanoi, Vietnam

<sup>f</sup>Department of Chemical Engineering, The University of Da Nang, University of Science and Technology, Da Nang City, Vietnam

<sup>g</sup>NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam.

E-mail: [ptthuong@ntt.edu.vn](mailto:ptthuong@ntt.edu.vn)

† Electronic supplementary information (ESI) available. See DOI: [10.1039/d1ra02529e](https://doi.org/10.1039/d1ra02529e)



(CPU) time consumption.<sup>23,24</sup> It is worthy to mention that the computational combination of molecular docking and the FPL approach was validated on available inhibitors of SARS-CoV-2 Mpro and showed good agreement between calculated binding free energies and experimental values.<sup>13,14,25</sup>

Computational drug repositioning has been a promising strategy to discover agents from known drugs for efficient response to the COVID-19 pandemic.<sup>26–28</sup> Recently, 17 approved drugs have been identified as inhibitors of SARS-CoV-2 Mpro *via* molecular docking and molecular dynamics simulations.<sup>29</sup> Artificial intelligence (AI) was also used to repurpose existing drugs.<sup>30</sup> However, the binding free energy between ligands and SARS-CoV-2 Mpro was not calculated in those previous works, probably leading to insufficient binding affinities for drug repurposing of compounds identified. Notably, molecular docking and the FPL simulations were previously demonstrated that they could adopt high accurate results compared with the respective experiments.<sup>14</sup> Therefore, in this study, the possible inhibitors of SARS-CoV-2 Mpro were screened from 2100 FDA-approved drugs using a combination of molecular docking and the FPL simulations. The results revealed that twenty-seven compounds could form large binding affinities to SARS-CoV-2 Mpro. These compounds thus become potential candidates for drugs against the COVID-19 disease effectively.

## Materials and methods

### Structure of SARS-CoV-2 Mpro and ligands

The crystal structure of SARS-CoV-2 Mpro in monomeric form (6Y2F) was downloaded from the Protein Data Bank.<sup>11</sup> The structure of ligands was obtained from the ZINC15 sub-database named FDA-approved drugs.<sup>31</sup>

### Molecular docking simulations

The Autodock Vina version 1.1 package was utilized to dock the screened ligands to the SARS-CoV-2 Mpro.<sup>32</sup> The parameters of the molecular docking simulations were prepared by Auto-dockTools 1.5.6 (ref. 33) following the earlier studies.<sup>34,35</sup> In detail, the exhaustiveness parameter of global searching that corresponds to the accuracy of the docking simulation was defined as 8 which represents default options. The atomic charges of protein and ligands were anticipated using the Gasteiger–Marsili approach.<sup>36,37</sup> The protein and ligands were emblemed by a united atom model with explicit polar hydrogens.<sup>38</sup> The maximum energy difference between the worst and best docking modes was set to 7 kcal mol<sup>−1</sup>. The grid center of Vina docking was selected as the center of mass of compound  $\alpha$ -ketoamide **13b**, which was obtained using the experimental pose.<sup>11</sup> The highest binding affinity profiles were chosen as the best docking conformation. The docking grid was determined as 2.6  $\times$  2.6  $\times$  2.6 nm which is able to accommodate the whole targeted active site.<sup>13–15</sup>

### Molecular dynamics simulations

We used the GROMACS version 5.1.5 (ref. 39) to simulate the conformation change of the complex SARS-CoV-2 Mpro and

ligand. The protein and ions were parameterized using the Amber99SBILDN force field.<sup>40</sup> The water molecules were topologized *via* the TIP3P water model.<sup>41</sup> The general Amber force field (GAFF)<sup>42</sup> was applied to represent the ligand using AmberTools18.<sup>43</sup> Combining the AMBER99SB-ILDN force field and water model provides one of the best options to estimate the free energy.<sup>44,45</sup> The restrained electrostatic potential (RESP) method<sup>46</sup> was used to fit the ligand atomic charges based on density functional theory (DFT) calculations using the combination of the B3LYP functional and 6-31G(d,p) basis set. The AMBER force field format was transformed to GROMACS using the ACPYPE protocol.<sup>47</sup> The time steps of MD simulations were set to 2 fs. MD simulations were carried out with an integrator at the absolute temperature of 300 K controlled by V-rescale. The 0.9 nm cutoff was applied to the non-bonded atoms pair list. The steepest descent method was initially applied to the complex of SARS-CoV-2 Mpro and inhibitor to minimize the system. To relax the system, the 0.1 ns of NVT and 2.0 ns of NPT simulations were computed. During the NVT and NPT ensembles, the C<sub>α</sub> atoms of SARS-CoV-2 Mpro were positionally restrained by a small harmonic force of 1000 kJ mol<sup>−1</sup> nm<sup>−2</sup>. The solvated complexes were recorded every 10 ps over MD simulations.

### Fast pulling of ligand simulations

The last snapshot of NPT simulations was then applied as the starting conformation of the steered-MD (SMD) simulation.<sup>23</sup> The complex of SARS-CoV-2 and ligand was covered into a rectangular periodic boundary conditions box (9.83  $\times$  5.92  $\times$  8.70 nm).<sup>13,25</sup> More than 50 000 atoms including SARS-CoV-2 Mpro,

**Table 1** The binding affinity values of verified SARS-CoV-2 Mpro inhibitors obtained *via* the docking simulations

| No. | Inhibitor       | $\Delta G_{\text{Dock}}^a$ | $\Delta G_{\text{exp}}^b$ |
|-----|-----------------|----------------------------|---------------------------|
| 1   | <b>11a</b>      | −9.96                      | −7.60                     |
| 2   | <b>11b</b>      | −10.13                     | −7.20                     |
| 3   | <b>11r</b>      | −9.23                      | −6.90                     |
| 4   | <b>13a</b>      | −7.70                      | −6.80                     |
| 5   | <b>13b</b>      | −8.45                      | −6.70                     |
| 6   | Carmofur        | −7.86                      | −6.60                     |
| 7   | Disulfiram      | −6.89                      | −6.10                     |
| 8   | Ebselen         | −8.45                      | −5.70                     |
| 9   | PX-12           | −6.39                      | −5.60                     |
| 10  | Shikonin        | −6.58                      | −3.90                     |
| 11  | Tideglusib      | −7.95                      | −3.80                     |
| 12  | Digitoxin       | −9.09                      | −8.00                     |
| 13  | Oubain          | −9.6                       | −7.20                     |
| 14  | Remdesivir      | −6.96                      | −6.40                     |
| 15  | Oxyclozanide    | −7.44                      | −6.30                     |
| 16  | Ebastine        | −7.06                      | −6.10                     |
| 17  | Toremifene      | −7.46                      | −5.90                     |
| 18  | Hexachlorophene | −8.28                      | −5.70                     |
| 19  | Chloroquine     | −6.74                      | −5.60                     |
| 20  | Triparanol      | −7.05                      | −5.50                     |
| 21  | Favipiravir     | −4.52                      | −4.80                     |

<sup>a</sup> The docking scores  $\Delta G_{\text{Dock}}$  were obtained *via* the Autodock Vina package. <sup>b</sup> The experimental binding free energy  $\Delta G_{\text{exp}}$  was roughly identified using the reported values<sup>50</sup> of inhibition constant  $IC_{50}$  with an assumption that  $IC_{50}$  is equal to  $k_i$ . The unit is in kcal mol<sup>−1</sup>.

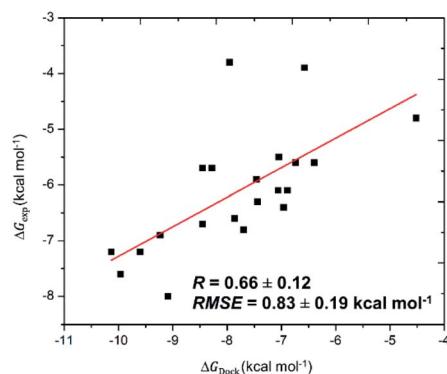


Fig. 1 Correlation between molecular docking and experiment. The calculated errors are the standard error.

ligand, water molecules and counter-balanced ions  $\text{Na}^+$  were involved in the simulated system. To pull the ligand out of the binding site in SARS-CoV-2 Mpro, an external harmonic force  $F = k(vt - z)$  along the  $Z$ -axis was applied on the center of mass of the ligand in the complex system (a spring constant of  $k = 0.005 \text{ nm ps}^{-1}$ , pulling speed of  $v = 600 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , and  $z$  is the displacement of the ligand mass center from its initial position).<sup>23,25,48</sup> The work of external force  $W$  is calculated according to the following equation:

$$W = v \int_0^t F(t) dt \quad (1)$$

Table 2 Twenty-seven top-lead compounds obtained by docking and FPL simulations

| No. | Name                        | ZINC ID          | Binding energy (kcal mol <sup>-1</sup> ) | Average force (pN) | Average work (kcal mol <sup>-1</sup> ) | Predicted $\Delta G_{\text{Pre}}^{\text{FPL}}$ |
|-----|-----------------------------|------------------|--|--------------------|--|--|
| 1   | Indocyanine green acid form | ZINC000008101127 | -8.6                                     | 861.75 ± 42.65     | 94.55 ± 5.76                           | -10.81   |
| 2   | Daclatasvir                 | ZINC000068204830 | -8.1                                     | 590.59 ± 39.59     | 85.57 ± 6.52                           | -10.30   |
| 3   | Teniposide                  | ZINC000004099009 | -8.4                                     | 636.29 ± 28.39     | 71.74 ± 3.17                           | -9.53  |
| 4   | Etoposide                   | ZINC000003938684 | -8.1                                     | 652.51 ± 37.03     | 67.02 ± 4.55                           | -9.27  |
| 5   | Levoleucovorin              | ZINC000009212427 | -8.1                                     | 610.87 ± 12.16     | 66.36 ± 1.89                           | -9.23  |
| 6   | Naldemedine                 | ZINC000100378061 | -8.2                                     | 648.26 ± 36.79     | 66.14 ± 3.12                           | -9.22  |
| 7   | Cabozantinib                | ZINC000070466416 | -8.1                                     | 621.70 ± 28.96     | 62.88 ± 2.86                           | -9.03  |
| 8   | Irinotecan                  | ZINC000001612996 | -9.0                                     | 625.12 ± 41.31     | 62.15 ± 4.92                           | -8.99  |
| 9   | Azilsartan medoxomil        | ZINC000014210642 | -8.2                                     | 550.68 ± 22.92     | 58.67 ± 3.46                           | -8.80  |
| 10  | Ergotamine                  | ZINC000052955754 | -8.4                                     | 576.51 ± 25.68     | 58.63 ± 2.61                           | -8.80  |
| 11  | Cromolyn                    | ZINC000253632968 | -8.1                                     | 551.18 ± 37.07     | 55.90 ± 5.61                           | -8.64  |
| 12  | Glecaprevir                 | ZINC000164528615 | -8.5                                     | 551.22 ± 40.48     | 55.54 ± 3.99                           | -8.62  |
| 13  | Dolutegravir                | ZINC000058581064 | -8.2                                     | 559.65 ± 24.05     | 55.05 ± 2.07                           | -8.59  |
| 14  | Saquinavir                  | ZINC000029416466 | -8.1                                     | 514.50 ± 31.70     | 54.97 ± 5.12                           | -8.59  |
| 15  | Dihydroergotamine           | ZINC000003978005 | -8.7                                     | 543.99 ± 42.88     | 54.42 ± 5.10                           | -8.56  |
| 16  | Accolate                    | ZINC00000896717  | -8.7                                     | 498.69 ± 32.54     | 52.91 ± 4.84                           | -8.47  |
| 17  | Lumacaftor                  | ZINC000064033452 | -8.5                                     | 528.13 ± 36.91     | 50.37 ± 3.87                           | -8.33  |
| 18  | Lifitegrast                 | ZINC000084668739 | -8.3                                     | 521.70 ± 22.03     | 49.47 ± 2.43                           | -8.28  |
| 19  | Doxazosin                   | ZINC000094566092 | -8.1                                     | 524.16 ± 17.71     | 48.82 ± 2.18                           | -8.25  |
| 20  | Rifaximin                   | ZINC000169621200 | -8.3                                     | 530.94 ± 42.43     | 45.23 ± 5.55                           | -8.04  |
| 21  | Ceftaroline                 | ZINC000003989268 | -8.2                                     | 420.93 ± 53.33     | 44.59 ± 5.96                           | -8.01  |
| 22  | Dutasteride                 | ZINC000003932831 | -8.1                                     | 499.62 ± 22.11     | 43.85 ± 2.42                           | -7.97  |
| 23  | Imatinib                    | ZINC000019632618 | -8.3                                     | 474.81 ± 16.15     | 43.26 ± 3.02                           | -7.93  |
| 24  | Raltegravir                 | ZINC00013831130  | -8.1                                     | 487.90 ± 22.66     | 43.05 ± 1.93                           | -7.92  |
| 25  | Trypan blue                 | ZINC000169289767 | -9.2                                     | 409.97 ± 45.04     | 41.79 ± 7.06                           | -7.85  |
| 26  | Nilotinib                   | ZINC000006716957 | -8.5                                     | 486.41 ± 26.65     | 39.98 ± 2.79                           | -7.75  |
| 27  | Regorafenib                 | ZINC000006745272 | -8.3                                     | 425.92 ± 16.90     | 37.54 ± 2.99                           | -7.61  |

During SMD simulations, the solvated complex systems were recorded the external pulling force and the ligand displacement every 0.1 ps for the estimation of the ligand-binding affinity.<sup>23</sup> For better sampling, 8 non-equilibrium MD simulation trajectories were independently conducted for each ligand–Mpro complex beginning with the same initial crystal structure but different random velocities. The mean external work  $\langle W \rangle$  was averaged from 8 SMD trajectories for each complex.

## Analyzed tools

To predict the ligand protonation state, the Chemicalize tools (<http://www.chemicalize.com>), a website application of the ChemAxon, were utilized. A hydrogen bond (HB) is defined if the angle of an acceptor (A)–hydrogen (H)–donor (D) is larger than  $135^\circ$  with the distance from A to D is smaller than 0.35 nm. A sidechain contact is determined if the distance between non-hydrogen atoms of SARS-CoV-2 Mpro and the ligands is smaller than 0.45 nm. The interaction between the SARS-CoV-2 Mpro protein and the ligand was illustrated by the molecular modeling software Maestro (free version).<sup>49</sup>

## Results and discussion

### Molecular docking simulation

The SARS-CoV-2 Mpro in monomeric form can be used as a target for CADD to prevent the function of SARS-CoV-2 Mpro<sup>13,14,25</sup> Autodock Vina,<sup>32</sup> an open-source docking package, is one of the most popular docking protocols to rapidly estimate



the binding affinity and binding pose of the protein and ligand complex. To validate the suitability of the approach, Autodock Vina was first applied to dock 21 SARS-CoV-2 Mpro inhibitors that have been confirmed experimentally. As was expected, the approach formed appropriate results for calculating the ligand-binding affinities of SARS-CoV-2 Mpro (Table 1) with a correlation coefficient of  $R_{\text{Dock}} = 0.66 \pm 0.12$  (Fig. 1), which is in good agreement with the recent work.<sup>14</sup> Additionally, the root-mean-square error (RMSE) concerning practical data was calculated as  $\text{RMSE} = 0.83 \pm 0.19 \text{ kcal mol}^{-1}$ . Noted that the results reported are consistent with the earlier studies.<sup>14</sup>

Based on the above results, the binding affinities between the SARS-CoV-2 Mpro and the screened ligands were first predicted *via* Autodock Vina. The docking results revealed that the calculated binding free energies for all computed complexes varied from  $-2.4$  to  $-9.2 \text{ kcal mol}^{-1}$ . The average of binding free energies obtained of  $-5.99 \text{ kcal mol}^{-1}$  with the standard error of the mean of  $1.08 \text{ kcal mol}^{-1}$ . Based on the results, 27 top-lead ligands having the binding energy with SARS-CoV-2 Mpro more negative than  $-8.1 \text{ kcal mol}^{-1}$  (Table 2) were chosen to be further examined by the FPL simulations.

## Docking binding pose

To investigate the binding interaction between the 27 top-lead compounds and SARS-CoV-2 Mpro, a detailed analysis of their possible docked conformation was performed to explore their interaction with the SARS-CoV-2 Mpro-binding pocket. Accordingly, the preferential binding pose of top-lead drugs obtained by docking simulations in the complex with SARS-CoV-2 Mpro was determined. The detailed interaction of SARS-CoV-2 Mpro and the full top-lead compounds are illustrated in Fig. S1.<sup>†</sup> In particular, Fig. 2 shows the particular binding of representative drugs including daclatasvir, teniposide, etoposide, and levoleucovorin with SARS-CoV-2 Mpro. Note that the substrate-binding site of SARS-CoV-2 Mpro resides in a cleft between domain I and domain II.<sup>11,51</sup> The obtained results suggest that the top-lead drugs identified by the docking method bind to the substrate-binding cleft of SARS-CoV-2 Mpro *via* different HBs. These compounds establish the sidechain interactions with a set of critical residues, including Thr26, His41, Leu141, Gly143, Ser144, Cys145, His163, Glu166, and Gln189, of SARS-CoV-2 Mpro. It should be noted that His41 and Cys145 reside

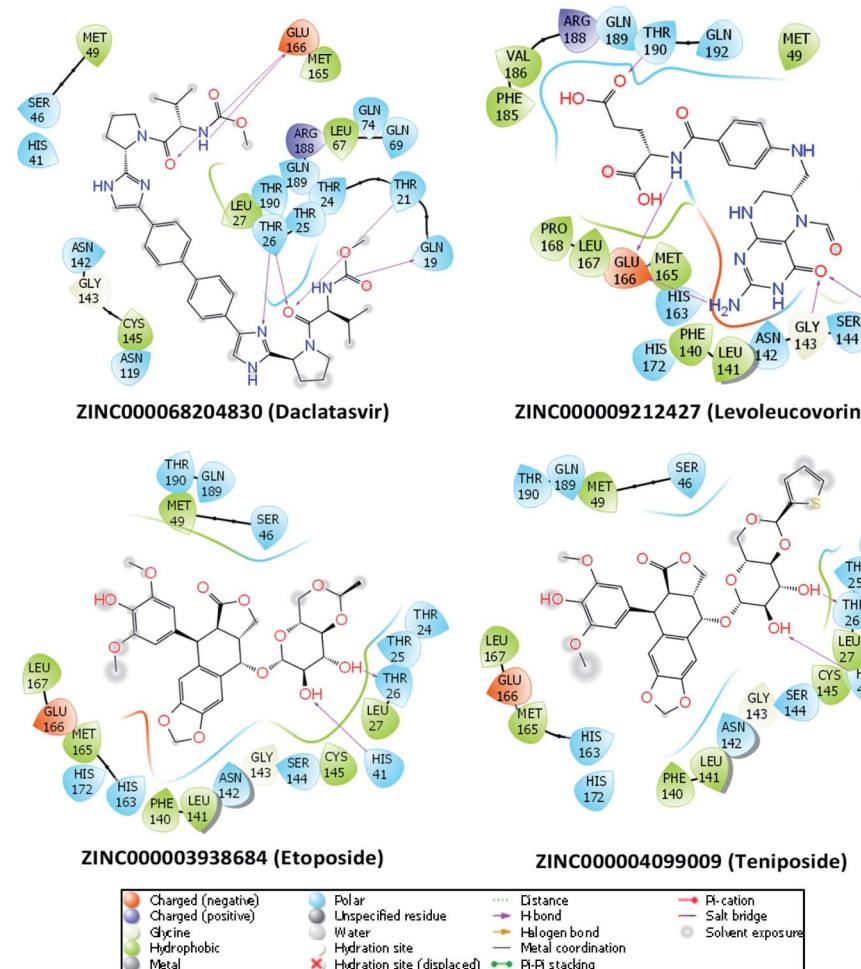


Fig. 2 The detailed interactions between SARS-CoV-2 Mpro and four drugs obtained via molecular docking simulations are illustrated by the molecular modeling software Maestro (free version).<sup>49</sup> HBs formed by residues of SARS-CoV-2 Mpro and ligands are indicated by purple arrows. Atoms of carbon, oxygen, nitrogen, and sulfur are presented in black, red, blue, and yellow, respectively.



in the Cys–His catalytic dyad and Glu166 is essential for dimerization of SARS-CoV-2 Mpro.<sup>7,11</sup>

### FPL simulation

The molecular docking simulations were performed with numerous limitations such as inflexible receptor and small number trial positions of inhibitors, the MD/SMD simulations were thus executed to refine the outcomes.<sup>52,53</sup> In this context, the FPL simulations were manipulated to re-rank the ligand-binding affinity to the SARS-CoV-2 Mpro since the approach successfully formed an appropriate outcome compared with the respective experiments,  $R = 0.76$ .<sup>14</sup> During MD/SMD calculations, the complexes were first relaxed to the equilibrated states. The clustering method was then employed to estimate the stabilized structures of the complexes of SARS-CoV-2 Mpro and ligands with an all-atom root-mean-square deviation (RMSD) cutoff of 3.0 Å. The dominant conformations of the top-lead compounds in the complex with SARS-CoV-2 Mpro are depicted in Fig. S2.† Since changing from the implicit solvent environment (docking results) to the explicit solvent modeling (MD simulations), the complex structures were slightly refined with an averaged RMSD between the initial and MD refine of ligands of *ca.* 2.0 Å. The structural changes of representative compounds are described in Fig. 3. Although the value of RMSD is small, implying the success of the docking calculation, some charged groups of the ligands were also rotated and then formed HBs to the receptor. The shifts are small but important, resulting in the difference of the affinity ranking order of ligands (*cf.* Table 2).

Over the FPL simulations, the maximal value of external force called rupture force and the pulling work averaged from 8 independent SMD simulation trajectories were both used as criteria to rank the ligand affinity. However, the pulling work is more appropriate than the rupture force because it directly

associates with the ligand-binding free energy *via* isobaric-isothermal Jarzynski's equality.<sup>23</sup> The average of pulling forces of all complexes between SARS-CoV-2 and tested ligands in a time-dependent manner is shown in Fig. S3.† The obtained results indicated that the mean rupture forces ranged from  $409.97 \pm 45.04$  (trypan blue free acid) to  $861.75 \pm 42.65$  (indocyanine green acid form) pN with the average value of  $552 \pm 31.4$  pN (Table 2). The pulling work was shown to be a key factor representing the binding of ligands and the protein.<sup>23,24</sup> According to FPL calculations, the average pulling works of the tested 27 compounds fell within a range of  $37.54 \pm 2.99$  (regorafenib) to  $94.55 \pm 5.76$  (indocyanine green acid form) kcal mol<sup>-1</sup> with the mean value of  $55.94 \pm 3.02$  kcal mol<sup>-1</sup> (Table 2).

Every complex of SARS-CoV-2 Mpro and ligand was computed over 8 independent FPL simulations which initiated from the same conformation but different in random velocity. One FPL trajectory consists of 0.1 ns of NVT, 2.0 ns of NPT, and 0.5 ns of SMD simulations. Totally, 20.8 ns of MD simulations were computed to evaluate the ligand-binding affinity with SARS-CoV-2 Mpro. The binding affinity of a ligand to the SARS-CoV-2 Mpro can thus be simulated 8 times during approximately 6 hours. Consequently, without the requirement of a professional computing system, the low CPU consumption allows the precise evaluation of the binding affinity of various compounds with SARS-CoV-2 Mpro at an appropriate time.

Recently, the FPL calculations have been proved to adopt a good agreement with the experimental data for SARS-CoV-2 Mpro and its inhibitors.<sup>13,14</sup> The estimated binding free energies  $\Delta G_{\text{Pre}}^{\text{FPL}}$  between top-lead ligands and SARS-CoV-2 Mpro were thus calculated in the same way as in the previous study.<sup>14</sup> The obtained results are shown in Table 2. The ligand with the predicted  $\Delta G_{\text{Pre}}^{\text{FPL}}$  more negative than  $-9.0$  kcal mol<sup>-1</sup> is thus strongly expected to be capable of inhibiting the function of the SARS-CoV-2 Mpro protein. Noted that indocyanine green acid form is a fluorescent dye used in medical diagnostics as an indicator substance in cardiac, circulatory, hepatic, and ophthalmic conditions<sup>54</sup> which might not be suitable to serve as a drug for COVID-19 treatment. Consequently, seven approved drugs including daclatasvir, teniposide, etoposide, levoleucovorin, naldemedine, cabozantinib, and irinotecan were predicted as the very promising inhibitors of SARS-CoV-2 Mpro in practice due to their high binding affinities calculated (Table 2).

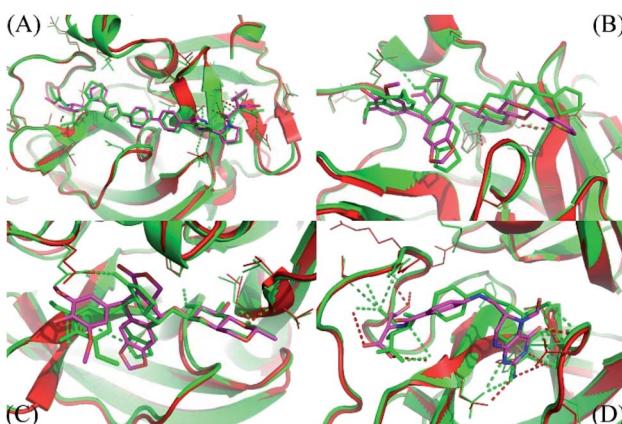


Fig. 3 The comparison between MD refined conformations of the complexes and docked structures. The MD refined structure was obtained by all-atom clustering with a cut-off of 0.3 nm over the last NPT snapshots. (A) Is the daclatasvir complex; (B) is the teniposide system; (C) is the etoposide complex; (D) is the levoleucovorin system. Both receptors and ligands obtained from docking are displayed in green. The figure was generated using PyMOL 1.3 open source.

### The most potential inhibitors of SARS-CoV-2 Mpro identified

The 2D structures of 7 top-lead ligands of SARS-CoV-2 Mpro ranked by the FPL simulations are illustrated in Fig. 4. Except for levoleucovorin and cabozantinib of which molecular structures are considered as flexible, the other 5 compounds possess the rigid structure and thus, lacking molecular flexibility. However, the substrate-binding cleft on the surface of SARS-CoV-2 Mpro is expected to be flexible to sizably accommodate a broad type of compound. Importantly, all of the 7 top-lead inhibitors predicted have many sites to form intermolecular HBs with the protein, hence being able to strongly interact with SARS-CoV-2 Mpro.



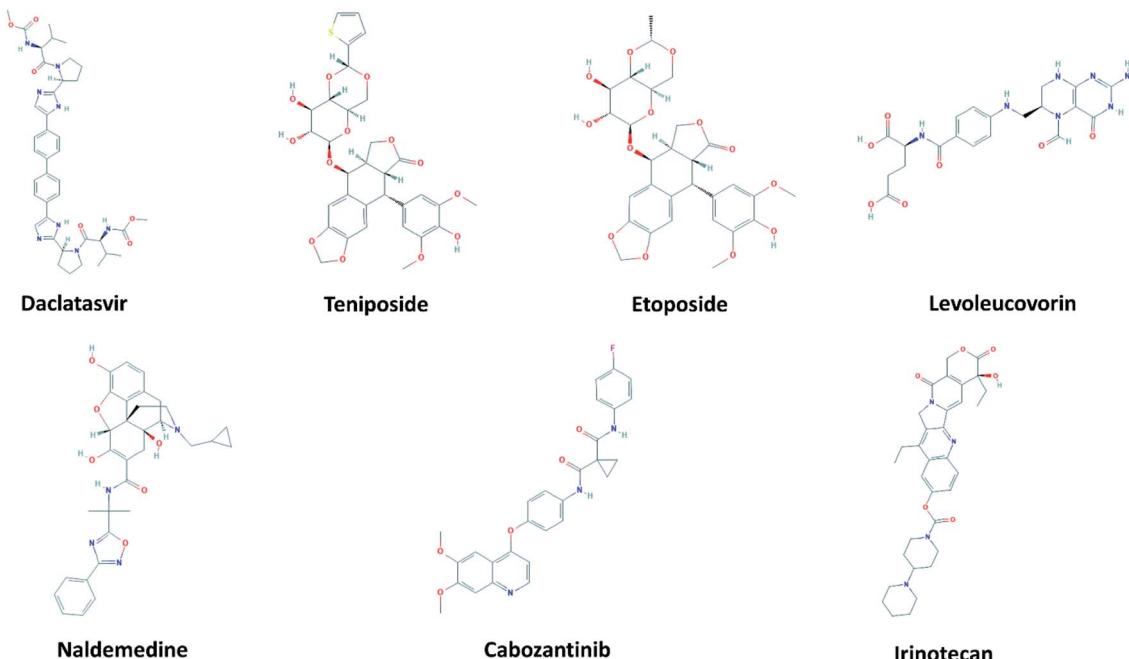


Fig. 4 2D structure of potential inhibitors for SARS-CoV-2 Mpro predicted by molecular docking and FPL simulations from the ZINC15 sub-database named FDA-approved drugs. The 2D structures were downloaded from an open chemistry database PubChem.<sup>60–66</sup>

Daclatasvir belonging to a valine and derivatives group is a direct-acting antiviral agent against Hepatitis C Virus (HCV), a positive-sense single-stranded RNA virus.<sup>55</sup> Previously, the anti-HCV drug was shown to be effective in the treatment of Middle East Respiratory Syndrome (MERS) coronavirus.<sup>56</sup> Therefore, anti-HCV drugs are expected to express a common antiviral activity against human coronaviruses. From our calculation, daclatasvir showed the highest affinity to SARS-CoV-2 Mpro among tested compounds (*cf.* Table 2) *via* forming HBs with a critical residue Glu166 and other residues, namely Thr21 and Thr26 (Fig. 2). Notably, it was shown that the combined use of sofosbuvir and daclatasvir improved the survival and clinical recovery of COVID-19 patients with modest to intense symptoms.<sup>57</sup> The introduction of sofosbuvir and daclatasvir to standard care could reduce the hospitalization time for COVID-19 patients in comparison to standard care alone.<sup>58,59</sup> These results indicated that daclatasvir is very promising in the treatment of COVID-19 patients and that our approach successfully selected the practically potential drug for COVID-19 therapy.

It is believed that both virus-infected and cancer cells require the elevation of nucleic acids and protein synthesis and energy metabolism. Thus, drugs inhibiting cancer cells can be effective in the suppression of viral replication. Indeed, different anti-neoplastic agents re-purposed for COVID-19 therapy have been applied for early clinical trials.<sup>67,68</sup> Etoposide and teniposide are anticancer drugs for the treatment of several types of tumors, leukemia, and lymphoma.<sup>69</sup> Etoposide and teniposide are both semisynthetic analogs of podophyllotoxin. They share a similar basic structure of the parent podophyllotoxin while the carbohydrate moiety of a methyl group in etoposide is substituted for

a phenylidene group in teniposide,<sup>69</sup> inducing a slight difference in binding affinities to SARS-CoV-2 Mpro (*cf.* Table 2). These two compounds are estimated to establish HBs with key residues Cys145 and His41 and residue Thr26 (Fig. 2). Teniposide possesses a log *P* value of 2.78 (DrugBank, Accession Number DB00444) while etoposide has a lower log *P* value of 1.16 (DrugBank, Accession Number DB00773). These values are in the optimal range of 1 to 3 for a compound to achieve appropriate physicochemical characteristics.<sup>70</sup> However, teniposide which is predicted to have a higher affinity to SARS-CoV-2 Mpro has better membrane permeability. Irinotecan is indicated for colorectal and pancreatic cancer treatment or combined with cisplatin for the cure of small cell lung cancer.<sup>71–73</sup> The Sn38 moiety is believed to mainly impact the high binding affinity of irinotecan to SARS-CoV-2 Mpro. The compound is predicted to form HBs with residues Thr24 and Thr26 of SARS-CoV-2 Mpro (Fig. S1†). Irinotecan (DrugBank, Accession Number DB00762) has a similar log *P* value and molecular weight to teniposide, indicating a comparable membrane permeability. In general, these three anti-cancer drugs are highly promising in the treatment of COVID-19.

Levoleucovorin, a folate analog, is utilized in rescue therapy to recover cells from the toxic effects of folate antagonists such as methotrexate after high-dose treatment in osteosarcoma therapy.<sup>74</sup> Naldemedine is an opioid receptor antagonist and used for the treatment of opioid-induced constipation.<sup>75</sup> Levoleucovorin and naldemedine share similar estimated binding affinities to SARS-CoV-2 Mpro (*cf.* Table 2). While naldemedine is predicted to form HBs with residues Thr24 and Ser46 (Fig. S1†), levoleucovorin is expected to establish several HBs with both critical residues Cys145 and Glu166 and other



residues including Leu141, Gly143, Ser144, and Thr190 (Fig. 2). Nevertheless, levoleucovorin is hydrophilic with a low  $\log P$  value of  $-2.8$  (DrugBank, Accession Number DB11596), pointing out that the compound is not suitable as a drug for COVID-19 treatment. Meanwhile, naldemedine shows a  $\log P$  value of  $2.43$ , indicating its good ability to be transported through a cellular membrane.

Cabozantinib suppresses metastasis and oncogenesis by inhibiting receptor tyrosine kinases.<sup>76</sup> Recently, several kinase inhibitors have been re-purposed for COVID-19 therapy.<sup>77,78</sup> The molecular structure of cabozantinib is not too bulky, flexible, and contains fluoro, an element that can be found in many bioactive compounds. Cabozantinib is predicted to establish HBs with residues Ser46, Gly143 and a key residue Glu166 of SARS-CoV-2 Mpro (Fig. S1†). The high  $\log P$  value  $4.66$  of cabozantinib (DrugBank, Accession Number DB08875) indicating the considerable level of toxicity may limit the application of this drug in practical use.

## Conclusion

Since the COVID-19 pandemic spread throughout the world, several inhibitors of SARS-CoV-2 Mpro have been identified experimentally.<sup>11,79</sup> Based on CADD, various computational studies have also been conducted to search for the promising inhibitors of Mpro and other critical enzymes of SARS-CoV-2.<sup>13,14,80-84</sup> Previously, a combination of molecular docking and FPL simulations was proved to efficiently predict the binding affinity of a ligand to SARS-CoV-2 Mpro. In this study, the same approach was employed to estimate the promising inhibitors for SARS-CoV-2 Mpro from a set of 2100 FDA-approved drugs. The binding conformation of the top-lead compounds identified with SARS-CoV-2 Mpro was also analyzed. The detailed interactions of the seven top-lead drugs including daclatasvir, teniposide, etoposide, levoleucovorin, naldemedine, cabozantinib, and irinotecan that have the predicted binding free energies with SARS-CoV-2 Mpro less than  $-9.00$  kcal mol $^{-1}$  indicated that these drugs all occupied the substrate-binding pocket of SARS-CoV-2 Mpro and thus potentially hindered the protease activity of the enzyme. These drugs interact with important residues of SARS-CoV-2 Mpro including Thr26, His41, Leu141, Gly143, Ser144, Cys145, His163, Glu166, and Gln189. Further *in vitro* and *in vivo* investigations are needed to be performed to validate the obtained results.

## Conflicts of interest

The authors declared that they have no conflicts of interest to this work.

## References

- 1 D. Cucinotta and M. Vanelli, *Acta Biomed.*, 2020, **91**, 157–160.
- 2 Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau and Z. Feng, *N. Engl. J. Med.*, 2020, **382**, 1199–1207.
- 3 K. Mizumoto and G. Chowell, *Emerging Infect. Dis.*, 2020, **26**, 1251–1256.
- 4 Worldometers, *COVID-19 Corona Virus Pandemic*, accessed 3rd May, 2021, <http://www.worldometers.info/coronavirus/>.
- 5 F. He, Y. Deng and W. Li, *J. Med. Virol.*, 2020, **92**, 719–725.
- 6 A. A. T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I. K. Singh, A. Singh, S. M. Atif, G. Hariprasad, G. M. Hasan and M. I. Hassan, *Biochim. Biophys. Acta, Mol. Basis Dis.*, 2020, **1866**, 165878.
- 7 Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang and C. Peng, *Nature*, 2020, 1–5.
- 8 O. O. Olubiyi, M. Olagunju, M. Keutmann, J. Loschwitz and B. Strodel, *Molecules*, 2020, **25**, 3193.
- 9 T. Pillaiyar, M. Manickam, V. Namasivayam, Y. Hayashi and S.-H. Jung, *J. Med. Chem.*, 2016, **59**, 6595–6628.
- 10 S. Ullrich and C. Nitsche, *Bioorg. Med. Chem. Lett.*, 2020, 127377.
- 11 L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, *Science*, 2020, **368**, 409–412.
- 12 W. Dai, B. Zhang, X.-M. Jiang, H. Su, J. Li, Y. Zhao, X. Xie, Z. Jin, J. Peng and F. Liu, *Science*, 2020, **368**, 1331–1335.
- 13 S. T. Ngo, N. Quynh Anh Pham, L. Thi Le, D.-H. Pham and V. V. Vu, *J. Chem. Inf. Model.*, 2020, **60**, 5771–5780.
- 14 M. Q. Pham, K. B. Vu, T. N. H. Pham, L. H. Tran, N. T. Tung, V. V. Vu, T. H. Nguyen and S. T. Ngo, *RSC Adv.*, 2020, **10**, 31991–31996.
- 15 S. T. Ngo, H. M. Nguyen, P. M. Quan, V. K. Truong, N. T. Tung and V. V. Vu, *RSC Adv.*, 2020, **10**, 40284–40290.
- 16 Z. Li, X. Li, Y.-Y. Huang, Y. Wu, R. Liu, L. Zhou, Y. Lin, D. Wu, L. Zhang and H. Liu, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 27381–27387.
- 17 A. K. Ghosh, M. Brindisi, D. Shahabi, M. E. Chapman and A. D. Mesecar, *ChemMedChem*, 2020, **15**, 907–932.
- 18 D. Gentile, V. Patamia, A. Scala, M. T. Sciortino, A. Piperno and A. Rescifina, *Mar. Drugs*, 2020, **18**, 225.
- 19 W. Yu and A. D. MacKerell, in *Antibiotics*, Springer, 2017, pp. 85–106.
- 20 G. R. Marshall, *Annu. Rev. Pharmacol. Toxicol.*, 1987, **27**, 193–213.
- 21 N. Homeyer, F. Stoll, A. Hillisch and H. Gohlke, *J. Chem. Theory Comput.*, 2014, **10**, 3331–3344.
- 22 S. T. Ngo, K. B. Vu, L. M. Bui and V. V. Vu, *ACS Omega*, 2019, **4**, 3887–3893.
- 23 S. T. Ngo, H. M. Hung and M. T. Nguyen, *J. Comput. Chem.*, 2016, **37**, 2734–2742.
- 24 N. T. Mai, N. T. Lan, T. Y. Vu, P. T. M. Duong, N. T. Tung and H. T. T. Phung, *J. Mol. Graphics Modell.*, 2020, **100**, 107648.
- 25 N. M. Tam, P. C. Nam, D. T. Quang, N. T. Tung, V. V. Vu and S. T. Ngo, *RSC Adv.*, 2021, **11**, 2926–2934.
- 26 R. Alexpandi, J. F. De Mesquita, S. K. Pandian and A. V. Ravi, *Front. Microbiol.*, 2020, **11**, 1796.
- 27 W. R. Ferraz, R. A. Gomes, A. L. S. Novaes and G. H. Goulart Trossini, *Future Med. Chem.*, 2020, **12**, 1815–1828.
- 28 M. Hagar, H. A. Ahmed, G. Aljohani and O. A. Alhaddad, *Int. J. Mol. Sci.*, 2020, **21**, 3922.



29 M. M. Ghahremanpour, J. Tirado-Rives, M. Deshmukh, J. A. Ippolito, C.-H. Zhang, I. Cabeza de Vaca, M.-E. Liosi, K. S. Anderson and W. L. Jorgensen, *ACS Med. Chem. Lett.*, 2020, **11**, 2526–2533.

30 K. Gao, D. D. Nguyen, J. Chen, R. Wang and G.-W. Wei, *J. Phys. Chem. Lett.*, 2020, **11**, 5373–5382.

31 T. Sterling and J. J. Irwin, *J. Chem. Inf. Model.*, 2015, **55**, 2324–2337.

32 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.

33 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.

34 N. T. Nguyen, T. H. Nguyen, T. N. H. Pham, N. T. Huy, M. V. Bay, M. Q. Pham, P. C. Nam, V. V. Vu and S. T. Ngo, *J. Chem. Inf. Model.*, 2019, **60**, 204–211.

35 S. T. Ngo, T. H. Nguyen, N. T. Tung, P. C. Nam, K. B. Vu and V. V. Vu, *J. Comput. Chem.*, 2020, **41**, 611–618.

36 J. Gasteiger and M. Marsili, *Tetrahedron Lett.*, 1978, **19**, 3181–3184.

37 J. Gasteiger and M. Marsili, *Tetrahedron*, 1980, **36**, 3219–3228.

38 S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell and A. J. Olson, *Nat. Protoc.*, 2016, **11**, 905–919.

39 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19–25.

40 A. E. Aliev, M. Kulke, H. S. Khaneja, V. Chudasama, T. D. Sheppard and R. M. Lanigan, *Proteins: Struct., Funct., Bioinf.*, 2014, **82**, 195–215.

41 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

42 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, **25**, 247–260.

43 D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, *et al.*, *AMBER 18*, University of California, San Francisco, 2018.

44 H. Zhang, C. Yin, Y. Jiang and D. van der Spoel, *J. Chem. Inf. Model.*, 2018, **58**, 1037–1052.

45 H. Zhang, Y. Jiang, Z. Cui and C. Yin, *J. Chem. Inf. Model.*, 2018, **58**, 1669–1681.

46 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

47 A. W. S. Da Silva and W. F. Vranken, *BMC Res. Notes*, 2012, **5**, 1–8.

48 S. T. Ngo, M. T. Nguyen and M. T. Nguyen, *Chem. Phys. Lett.*, 2017, **676**, 12–17.

49 *Schrödinger Release 2021-1*, Maestro, Schrödinger, LLC, New York, NY, USA, 2021, <http://www.schrodinger.com/products/maestro>.

50 S. A. Amin, S. Banerjee, K. Ghosh, S. Gayen and T. Jha, *Bioorg. Med. Chem.*, 2020, **19**, 115860.

51 K. Anand, G. J. Palm, J. R. Mesters, S. G. Siddell, J. Ziebuhr and R. Hilgenfeld, *EMBO J.*, 2002, **21**, 3213–3224.

52 S. Decherchi and A. Cavalli, *Chem. Rev.*, 2020, **120**, 12788–12833.

53 S. Genheden and U. Ryde, *Expert Opin. Drug Discovery*, 2015, **10**, 449–461.

54 J. T. Alander, I. Kaartinen, A. Laakso, T. Pätilä, T. Spillmann, V. V. Tuchin, M. Venermo and P. Välisuo, *Int. J. Biomed. Imaging*, 2012, **2012**, 940585.

55 M. A. Smith, R. E. Regal and R. A. Mohammad, *Ann. Pharmacother.*, 2016, **50**, 39–46.

56 A. A. Elfiky, S. M. Mahdy and W. M. Elshemey, *J. Med. Virol.*, 2017, **89**, 1040–1047.

57 B. Simmons, H. Wentzel, S. Mobarak, G. Eslami, A. Sadeghi, A. A. Asgari, H. A. Kasgari, H. T. Fakheri, S. Merat and A. Hill, *J. Antimicrob. Chemother.*, 2020, **76**, 286–291.

58 G. Eslami, S. Mousaviasl, E. Radmanesh, S. Jelvay, S. Bitaraf, B. Simmons, H. Wentzel, A. Hill, A. Sadeghi and J. Freeman, *J. Antimicrob. Chemother.*, 2020, **75**, 3366–3372.

59 A. Sadeghi, A. Ali Asgari, A. Norouzi, Z. Kheiri, A. Anushirvani, M. Montazeri, H. Hosamirudsai, S. Afhami, E. Akbarpour and R. Aliannejad, *J. Antimicrob. Chemother.*, 2020, **75**, 3379–3385.

60 National Center for Biotechnology Information, *PubChem Compound Summary for CID 25154714*, *Daclatasvir*, accessed Apr. 26, 2021, <http://www.pubchem.ncbi.nlm.nih.gov/compound/25154714#section=2D-Structure>.

61 National Center for Biotechnology Information, *PubChem Compound Summary for CID 452548*, *Teniposide*, accessed Apr. 26, 2021, <http://www.pubchem.ncbi.nlm.nih.gov/compound/452548#section=2D-Structure>.

62 National Center for Biotechnology Information, *PubChem Compound Summary for CID 36462*, *Etoposide*, accessed Apr. 26, 2021, <http://www.pubchem.ncbi.nlm.nih.gov/compound/36462#section=2D-Structure>.

63 National Center for Biotechnology Information, *PubChem Compound Summary for CID 135398559*, *Levoleucovorin*, accessed Apr. 26, 2021, <http://www.pubchem.ncbi.nlm.nih.gov/compound/135398559#section=2D-Structure>.

64 National Center for Biotechnology Information, *PubChem Compound Summary for CID 54732242*, *Naldemedine*, accessed Apr. 26, 2021, <http://www.pubchem.ncbi.nlm.nih.gov/compound/54732242#section=2D-Structure>.

65 National Center for Biotechnology Information, *PubChem Compound Summary for CID 25102847*, *Cabozantinib*, accessed Apr. 26, 2021, <http://www.pubchem.ncbi.nlm.nih.gov/compound/25102847#section=2D-Structure>.

66 National Center for Biotechnology Information, *PubChem Compound Summary for CID 60838*, *Irinotecan*, accessed Apr. 26, 2021, <http://www.pubchem.ncbi.nlm.nih.gov/compound/60838#section=2D-Structure>.

67 G. Ciliberto, R. Mancini and M. G. Paggi, *J. Exp. Clin. Cancer Res.*, 2020, **39**, 1–9.

68 K. El Bairi, D. Trapani, A. Petrillo, C. Le Page, H. Zbakh, B. Daniele, R. Belbaraka, G. Curigliano and S. Afqir, *Eur. J. Cancer*, 2020, **141**, 40–61.



69 P. I. Clark and M. L. Slevin, *Clin. Pharmacokinet.*, 1987, **12**, 223–252.

70 J. A. Arnott and S. L. Planey, *Expert Opin. Drug Discovery*, 2012, **7**, 863–875.

71 M. Ramesh, P. Ahlawat and N. R. Srinivas, *Biomed. Chromatogr.*, 2010, **24**, 104–123.

72 K. Noda, Y. Nishiwaki, M. Kawahara, S. Negoro, T. Sugiura, A. Yokoyama, M. Fukuoka, K. Mori, K. Watanabe and T. Tamura, *N. Engl. J. Med.*, 2002, **346**, 85–91.

73 F. A. u. Rahman, S. Ali and M. W. Saif, *Ther. Adv. Gastroenterol.*, 2017, **10**, 563–572.

74 V. T. G. Chuang and M. Suno, *Ann. Pharmacother.*, 2012, **46**, 1349–1357.

75 H. A. Blair, *Drugs*, 2019, **79**, 1241–1247.

76 C. Grüllich, in *Small Molecules in Oncology*, Springer, 2014, pp. 207–214.

77 E. Weisberg, A. Parent, P. L. Yang, M. Sattler, Q. Liu, Q. Liu, J. Wang, C. Meng, S. J. Buhrlage and N. Gray, *Pharm. Res.*, 2020, **37**, 1–29.

78 S. Weston, C. M. Coleman, R. Haupt, J. Logue, K. Matthews, Y. Li, H. M. Reyes, S. R. Weiss and M. B. Frieman, *J. Virol.*, 2020, **94**, e01218–e01220.

79 T. Mohammad, A. Shamsi, S. Anwar, M. Umair, A. Hussain, M. T. Rehman, M. F. AlAjmi, A. Islam and M. I. Hassan, *Virus Res.*, 2020, **288**, 198102.

80 A. Romeo, F. Iacobelli and M. Falconi, *Virus Res.*, 2020, **286**, 198068.

81 D. Bhowmik, R. Nandi, R. Jagadeesan, N. Kumar, A. Prakash and D. Kumar, *Infect. Genet. Evol.*, 2020, **84**, 104451.

82 D. Dey, S. Borkotoky and M. Banerjee, *Comput. Biol. Med.*, 2020, **127**, 104063.

83 R. Chemboli, R. Kapavarapu, K. Deepti, K. Prasad, A. G. Reddy, A. N. Kumar, M. V. B. Rao and M. Pal, *J. Mol. Struct.*, 2021, **1230**, 129868.

84 S. Shahinshavali, K. A. Hossain, A. V. D. N. Kumar, A. G. Reddy, D. Kolli, A. Nakhi, M. V. B. Rao and M. Pal, *Tetrahedron Lett.*, 2020, **61**, 152336.

